Invited Paper

Internet Color Imaging

Hsien-Che Lee

Imaging Science and Technology Laboratory
Eastman Kodak Company, Rochester, New York 14650-1816 USA

ABSTRACT

The sharing and exchange of color images over the Internet pose very challenging problems to color science and technology. Emerging color standards will solve many of the problems we face today, but existing images of unknown origin and output devices of unknown calibration will continue to cause problems for many users. This paper presents a brief overview of available solutions to some of the problems and suggests some directions for future research. Although most existing solutions are quite primitive and fragile, the rapid advance of computing technology promises to bring more sophisticated and intelligent image processing algorithms to common practical use. Image understanding, scene physics, visual calibration, and image perception are four areas of research that are beginning to make good progress toward a fully automatic quality optimization system for color imaging applications.

Keywords: Internet, color imaging, automatic calibration, scene statistics, digital image processing

1. INTRODUCTION

When we order a sweater from a web site, how do we know if the color of the sweater is what we like? When we send a digital camera image to an on-line fulfillment center, how do we know if they can print it with good tone and color rendition? When we download a color image from a web site and print it on the color printer at our home, how can we make it come out as beautiful as we see it on our color monitor? These are questions that we are facing everyday. Color imaging applications for Internet shopping, services, and information gathering have become ubiquitous. Yet, color images that are exchanged over the Internet originate from widely varied sources, display/printing devices used to show those images are not calibrated, and viewing conditions are rarely controlled. So how can the whole thing work? Well, chaotic as it may be, there are several factors that save us from a total breakdown in such a mess: (1) Our visual system is very capable and very forgiving. With an amazing grace, it can often adjust to the distortion and extract the information needed. It is not that we do not see the distortion, it is that we choose not to pay too much attention to it. (2) Devices are built to vaguely conform to various standards, and those standards are not drastically different from each other. (3) We don't know what we are missing. Occasionally we see great pictures on our monitors and we are pleasantly surprised. We rarely ask, why don't we get great pictures all the time?

There are three basic classes of technical problems in Internet color imaging: (1) color images of unknown calibration, (2) imaging devices of unknown characteristics, and (3) viewing conditions of unknown perceptual effects. Solutions to each of these problems vary from completely manual to fully automatic adjustments, from closed systems to standardized interfaces, and from approximation to perfection. These problems and their possible solutions are discussed in this paper, and future research directions are suggested in the discussion.

2. STANDARDIZATION

The major component in the solution of most problems in Internet color imaging is to standardize the protocols of how color information should be communicated. The protocols have to be complete in all technical details. For example, it is not sufficient to specify the RGB signals from a digital camera as gamma-corrected video signals. Ideally, the spectral response functions of the camera should be provided with the camera. However, most consumers do not know how to take advantages of this type of technical information, or are unwilling to spend the time to do so. Therefore, national and international standardization efforts are mostly directed toward simple and packaged solutions. So instead of asking for the manufacturers to provide the technical information with the products, standards tend to describe a recommended system and its signal specifications, and it is up to

In *Input/Output and Imaging Technologies II*, Yung-Sheng Liu, Thomas S. Huang, Editors,
Proceedings of SPIE Vol. 4080 (2000) • 0277-786X/00/$15.00

each manufacturer to produce products that can work "reasonably well" with the standard system. This is a very practical and inexpensive solution, although it means that the needed technical information is often not available to the knowledgeable consumers. For example, chromaticity coordinates of the phosphors of a CRT monitor are usually not provided to the user.

Among the various international standard bodies, ISO, IEC, CIE, and ITU are four of the major organizations that publish standards of direct interest to the field of color imaging. The International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC), and the International Commission on Illumination (CIE) are the three major organizations that establish voluntary international standards. The International Telecommunication Union (ITU) is organized by the United Nations and its standards are regulatory through government administrations and treaties.[1,2] These international standards are then, in turn, used by industries to set up proposals for other applications. For example, the ITU-R Recommendation BT.709 (Parameter Values for the HDTV Standards for Production and International Programme Exchange) and Recommendation BT.1200 (Target Standards for Digital Video Systems for the Studio and International Programme Exchange) are two international standards that are widely adopted and adapted in color imaging applications, such as KODAK PHOTOYCC Color Interchange Space[3] and sRGB[4] color encodings.

In order to facilitate the standardization of color management systems, a non-profit organization, called International Color Consortium (ICC), was established in 1993.[5] The basic idea is to provide a device profile for every imaging device so that color data produced by one device can be translated into a device-independent profile connection space (PCS), which can, in turn, be converted into the native color space of another device. The ICC profile format is described by the document published by the Consortium. Although the interpretation of the rendering intent of some profile parameters can be ambiguous,[6] the ICC profiles, if fully implemented by most imaging hardware and software companies, will be a giant step forward toward solving many (but not all) problems in Internet color imaging applications. However, for adjustable devices, such as monitors, scanners, and digital cameras, a fixed device profile is obviously not sufficient. For example, if the contrast or brightness knob of a monitor is adjusted by the user, the monitor device profile is no longer valid for the status of that monitor.

A major advantage of the device profile approach to color management is that color images can be saved in the native color space of the device without unnecessary quantization to an intermediate color space.[7] This is especially important for 8 bits/color/pixel images. A fundamental problem with color solutions based on standards is that the color images are at best colorimetrically or perceptually correct (remember, this is a great position to be in), but may be far from visually optimum in quality. This is partially due to the fact that standards are related to systems and devices, not individual images. It is also partially due to our lack of understanding in human perception.

3. IMAGES OF UNKNOWN CALIBRATION

Most color images existing on Internet do not have any calibration information associated with them. Similarly, many images sent to on-line fulfillment centers are not calibrated. How do we deal with such images?

3.1. Interactive Mode

If a human operator is engaged in processing such images, a good strategy is to first estimate its basic tone scale metric. Are the digital numbers proportional to linear luminance, log luminance, or video (gamma corrected) luminance in the scene? These three are the most often used metrics from CCD sensors, film, and video cameras. We can make the assumed transform (with some variations in parameters) and display the image on a calibrated monitor to see which of these transforms look best and take a bet. However, most color images are produced through some nonlinear tone scale curves. Therefore, there will be a lot of work to adjust the highlight and the shadow to make the image look right with our intended tone reproduction curve. Because many Internet images are meant to be viewed on CRT monitors, they are very likely to be in gamma-corrected video space (such as NTSC-RGB or sRGB). The next step is to extract the digital color values from what we think are the neutral (gray or white) objects and the skin areas. The neutral objects will help us to do a basic color balance. The skin areas will allow us to estimate the color matrix required to rotate the color axes to the ones we want to use in our device. This is easier to say than done. Although the unexposed skin area of a given race tend to have a certain reflectance value, the exposed skin areas tend to vary greatly in lightness. Table 1 shows some sample

Table 1. Sample measurements of skin color (forehead and cheek).

African	$L^* = 37.6 \pm 1.3$	$a^* = 6.9 \pm 1.4$	$b^* = 10.7 \pm 2.3$
Arabian	$L^* = 61.5 \pm 2.3$	$a^* = 5.6 \pm 1.1$	$b^* = 17.3 \pm 1.8$
Caucasian	$L^* = 66.3 \pm 2.8$	$a^* = 11.2 \pm 0.9$	$b^* = 12.3 \pm 1.8$
Japanese	$L^* = 60.7 \pm 4.37$	$a^* = 10.8 \pm 2.36$	$b^* = 17.1 \pm 2.19$
Vietnamese	$L^* = 65.1 \pm 3.1$	$a^* = 5.4 \pm 0.8$	$b^* = 15.4 \pm 1.1$

measurements of (forehead/cheek) skin colors.[8–10] From the spectral measurement data reported by Edwards and Duntley,[11,12] we have computed the tristimulus values of the skin color of various races. In general, the dominant wavelength of (unexposed) skin is relatively constant across races (at about 584 nm under D_{65} illuminant). The main difference is in the luminance factor (varies from 7% to 45%) and the excitation purity (varies from 17% to 33% under D_{65}). The effect of sun tan is to shift the dominant wavelength toward a longer wavelength by an amount on the order of 10 nm, while the excitation purity remains about the same. Knowledge of typical distributions of skin color only gives us some estimates of how much and which way a color correction should be given. In the interactive mode, an operator can look at the image on a calibrated monitor and make continuous adjustments as needed.

3.2. Automatic Mode

For many applications, the cost and throughput requirements cannot afford too much operator intervention. In these cases, automatic algorithms have to be developed so that tone, color, and sharpness correction can be carried out all by computers, on an image by image basis. Automatic correction requires that the image calibration be estimated first and then the desired manipulation applied. The first step is the more difficult one of the two and currently, there is no good solution. However, there are potential research directions that we can see from some of the existing approaches.

In general, it is not possible to derive the exact relationships between the scene radiances and the digital values in a given digital image. However, it is interesting to note that when we display or print an image with a wrong calibration, we can often see from the rendered image that something is not quite "right". How can we sense that? What is it in the rendered image that is telling us that something is not right? Are there some "invariant" or "inherent" features in natural scenes that our visual system has learned to recognize and when those features are not reproduced well in an image because of wrong calibration, our visual system will sense the distortion of those features? It is difficult to imagine that such "invariant" features can exist. However, several studies have shown that indeed some characteristic features do exist for natural scenes. For example, the amplitude A of the radial spatial frequency f of natural scenes tend to be a power function[13–15] of the frequency f: $A(f) = \alpha f^{-p}$ and typical value of p is between 0.8 and 1.2. Because this characteristics is found to be relatively insensitive to calibration,[15] it is not useful for estimating the calibration from an image.

3.2.1. Tone correction

One of the features that has been proposed[16] for estimating the unknown tone scale calibration is the statistical property that the log-exposure distribution of a natural scene is approximately a Gaussian distribution. It is argued[16,17] that this property is due to the random distributions of surface orientation, reflectance factors, textures, and illumination, and also partially due to the central limit theorem. One interesting observation from the underlying heuristic reasoning is that the theoretical distribution holds true for any spectral response function used to measure exposure. This can be used for color correction for images that have mixed illuminants.[16] However, it is very easy to give counter examples in which such a Gaussian property does not hold for individual images or even for an ensemble of images,[18] depending on the contents on the images. For example, if we take an outdoor picture that includes a significant portion of the sky, the log-exposure distribution of the image is most likely to be bimodal. One might still argue that each mode of the histogram can be approximated by a Gaussian distribution. Unfortunately, even a mixture of Gaussians is often not a good model, because if there are one or more large uniform areas in the image, the shape of the log-exposure distribution will be quite varied. One way

to reduce the bias introduced by a large uniform area is to sample only where color or exposure signals change significantly.[16] The other way is to allow deviation from normality in a parameterized family of distributions.[19]

The problem of estimating the unknown calibration can be greatly simplified if we are interested in classifying the unknown input into one of the three most widely used metrics: linear-exposure, log-exposure, and video gamma-corrected exposure. For example, a simple way to classify images of unknown calibration is to take advantage of the fact that the histogram of a log-exposure image is more symmetric with respect to its mean than a linear-exposure image. If the histogram of the image in question is highly skewed to the right, it is more likely to be a linear-exposure image. The skewness of a distribution $h(x)$ can be measured by:

$$\text{skewness} = \frac{m_3}{\sigma^3}$$

where m_3 is the third central moment, i.e., $m_3 = \int (x-\mu)^3 \cdot h(x)\text{dx}$, and μ and σ are the mean and the standard deviation. We have calculated the skewness of the linear-exposure histograms and that of the log-exposure histograms for 1800 consumer images. Figure 1 shows the skewness distributions for these two image metrics. The two distributions intersect at skewness = 0.75. The fraction of log-exposure images that have a skewness

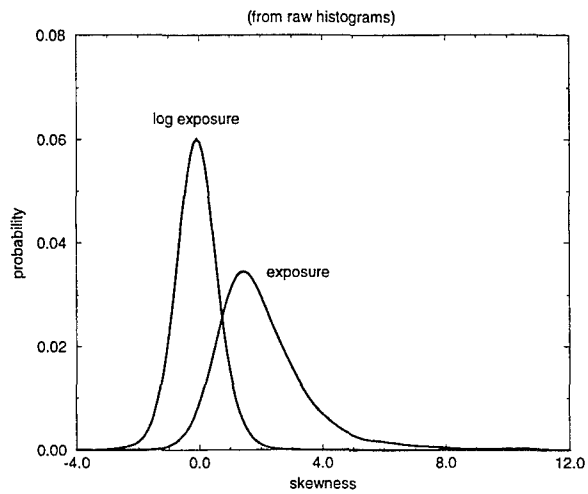

Figure 1. The distributions of skewness for exposure and log-exposure histograms for 1800 consumer images. (The two distribution curves have been smoothed.)

greater than 0.75 is about 11.2%. The fraction of linear-exposure images that have a skewness less than 0.75 is about 16.3%. Thus, from the skewness of the histogram alone, we can classify the input image into one of the two metrics (linear-exposure or log-exposure) correctly more than 80% of the time. In fact, we can improve the classification by using the skewness of the histogram accumulated only from the edge pixels, thus excluding the biases from large uniform areas. If we only sample along edge pixels in the images and calculate the skewness of the log exposure histograms of the edge pixels, we find that the standard deviation of the skewness distribution is now reduced from 0.59 to 0.42. The mis-classification rate of rejecting a log-exposure image has dropped to 3.8%. However, the above experimental results are based on the two-class discrimination problem. The algorithm does not work well when we have to deal with the three-class problem for linear, log, and video metrics.

Suppose that we have successfully estimated the unknown metrics of the input image, what can we do to improve the tone rendition of the image? This is a much easier problem. Although really robust algorithms are still being developed, several methods exist for adjusting global or local contrast of an image. For example, a global contrast adjustment algorithm[19] can take advantage of pre-compiled statistical data for some scene contrast estimator, such as the standard deviation, Δ, of the histogram of edge pixels. We have compiled such statistics. Figure 2 shows the distribution of Δ calculated for 1800 consumer images. Its mean is 0.375 in log-exposure. If we take $\pm$ 3 Δ (i.e. 6 standard deviations of the log-exposure histogram) as the dynamic luminance range of

the scene, we have an average log scene luminance range of $0.375 \times 6 = 2.25$, which corresponds to a luminance range of 168:1. This number is indeed very close to the average scene luminance range of 160:1 reported by Jones and Condit in their classical study using actual measurements on many natural scenes.[20] Another study[21] also reported that the average standard deviation of the log-exposure histograms is 0.33. The mutual confirmation of these studies does not mean that the current contrast estimate is accurate, but it shows that the algorithm can produce a reasonable result with a very simple computational procedure that does not require much prior knowledge. Further experimental testing is needed to achieve the optimal contrast adjustment. From Fig. 2 one

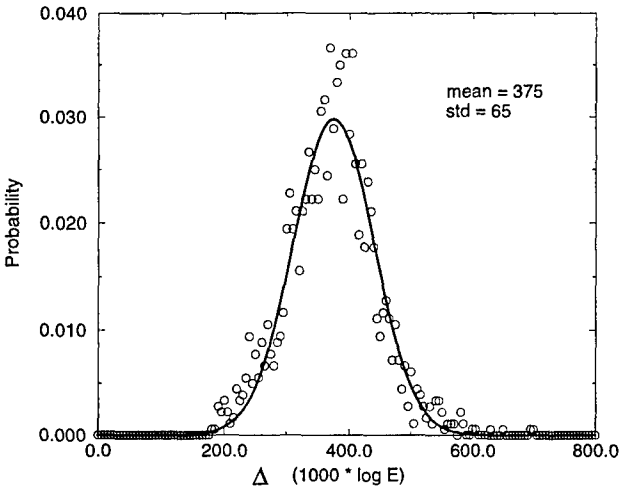

Figure 2. The distribution of the standard deviation, Δ, of the log-exposure histogram of the edge pixels of an image for 1800 consumer images.

can see that a scene dynamic range can be as high as $0.55 \times 6 = 3.3$ in log-exposure (or about 2000:1 in exposure) and as low as $0.2 \times 6 = 1.2$ in log-exposure (or about 16:1 in exposure). For most images of small dynamic range, say less than 80:1, experimental results so far show that the contrast adjustment greatly improves their perceived image quality.

In addition to contrast adjustment, it is also necessary to decide how light or how dark an image should be display or printed. This problem is called the density balance problem in photofinishing and it is similar to the exposure control problem in digital camera design. The problem is stated as follows: given a digital image, determine the digital value that is to be displayed at a given luminance level or printed at a given reflectance, so that the rendered image looks optimal in tonal quality. Most existing algorithms for density balance are proprietary and not available in public domain. A well-known algorithm is the integration-to-gray method[22] and its many variations. Using a consumer image database in which the optimum balance point (aim) for every image was determined by three experts, we have tested how well the simple integration-to-gray method works on consumer images. The database consists of 2697 images collected from consumers. The integration is done in two ways: averaging in exposure and averaging in log-exposure. The integrated red, green, blue values are used to compute a neutral balance point by the following equation:

$$L = \frac{1}{\sqrt{3}}(\log R + \log G + \log B). \tag{1}$$

This computed balance point is then compared with the experts' optimum point. Figure 3 shows the error distributions along this neutral "luminance" axis. There are two interesting observations from these statistical data: (a) Averaging in exposure and averaging in log-exposure yield about the same magnitude of estimation error. (b) Averaging in exposure and averaging in log-exposure have the opposite bias - the averaged exposure is lower, while the averaged log-exposure is higher than the experts' aim. In general, the gray world assumption produces a mediocre estimate for density balance. One of the obvious problem of the integration-to-gray method is that if there are large uniform areas in an image, they will bias the average to the luminances of those areas.

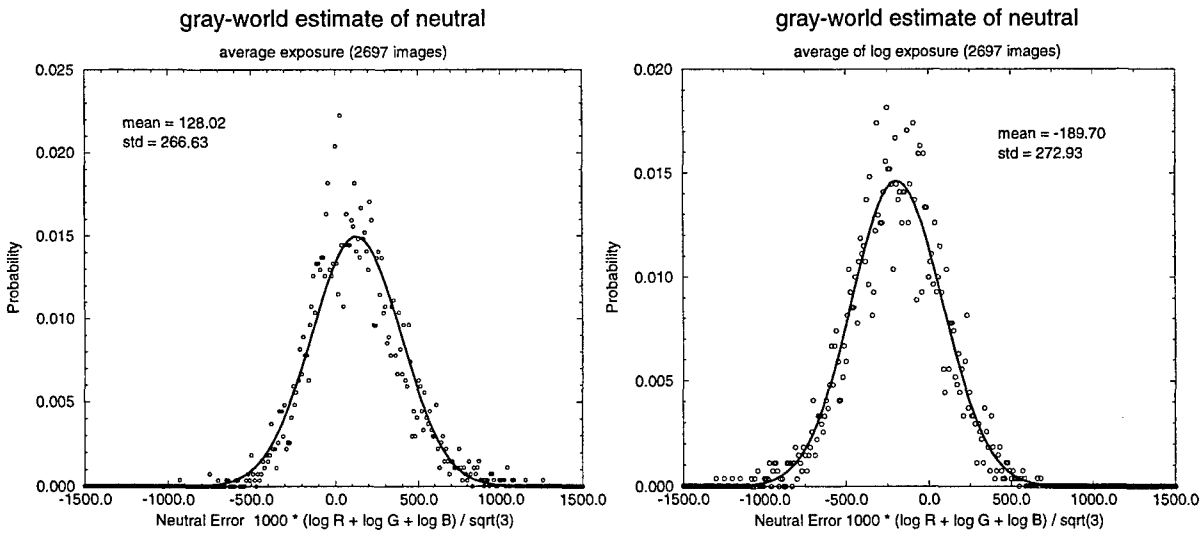

Figure 3. Comparison of Neutral error distributions. Left: averaging in exposure; Right: averaging in log-exposure.

Again, as we discussed before, an obvious improvement is to sample only on edge pixels[16,23] or from active (busy) image regions.[24,25]

3.2.2. Color correction

Similar to tone correction, there are two steps in color correction. The first step is to estimate the unknown color calibration and the second step is to apply the desired color manipulation (such as color balance) and enhancement (such as boosting color chroma).

In tone correction, the estimation of unknown calibration is mainly for deriving the functional relationship between the scene radiance and the digital image value. Usually, there is no explicit attempt to estimate how to combine the red, green, blue image values to produce what would be measured by the CIE luminous efficiency function $V(\lambda)$. The reason is that in tone correction, we are mainly interested in the intensity relationship between exposure and image value, rather than spectral relationship. When we deal with color correction, the spectral relations become important. It is no longer sufficient for us to know that our image values are proportional to linear-exposure or log-exposure. We actually have to know how they are related to the colors we see. Let R, G, B be the red, green, blue exposures of a pixel in an image and let X, Y, Z be the tristimulus values of the object color corresponding to that pixel. The simplest approximation of the relationship between R, G, B and X, Y, Z is a 3×3 matrix, M, i.e.,

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathrm{M} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{2}$$

Theoretically, a 3×3 matrix is an exact transformation if the spectral response functions of the image capture system are linear combinations of the CIE color matching functions. Because most imaging systems are far from that, a 3×3 matrix may not be a good approximation for our images at all. However, currently, this simple approximation is as complicated as we can try to estimate.

There are nine unknown elements in M to be estimated. Since in the image capture and printing processes, an overall scale factor can be and will be adjusted on an image by image basis. This is solved as the density balance problem we discussed in tone correction. The remaining eight unknowns can be determined from four pairs of corresponding chromaticity coordinates in (R, G, B) and (X, Y, Z). So, which four possible pairs can we estimate from an image automatically? Two important pairs are the neutral (gray) color and the skin color. The problem of estimating the neutral color in the image is called the color balance problem. The existing algorithms for solving the problem have been reviewed elsewhere.[26] Despite many new algorithms developed for

color constancy, the gray world assumption continues to be the backbone of the color correction algorithms for most printers and video cameras. But, just how gray is the world? If we average the exposure of all the pixels in a color image, we obtain 3 numbers: the average red, green, and blue values, which can be represented as a point in the three-dimensional (R,G,B) color space. In order to remove the exposure differences among images, the R,G,B aims (established by expert judges) of that image are subtracted from the image averages, so that if the averages predict the aims perfectly, the point representing the image should fall at the origin of the (R,G,B) color space. If we do this for 2697 images, we obtain a cluster of points, each representing an image. In order to show the error distribution from the gray world assumption, we project the errors to the red-blue direction and the magenta-green direction, because they are close to the eigenvectors of the covariance matrix computed from all the pixels in the 2697 image. The two chromatic axes are defined as:

$$s = \frac{(\log R - \log B)}{\sqrt{2}} \quad [\text{red} - \text{blue}]$$

$$t = \frac{(\log R - 2\log G + \log B)}{\sqrt{6}} \quad [\text{magenta} - \text{green}]$$

Figure 4 shows how the errors are distributed. As can be seen from these figures, the error distributions tend

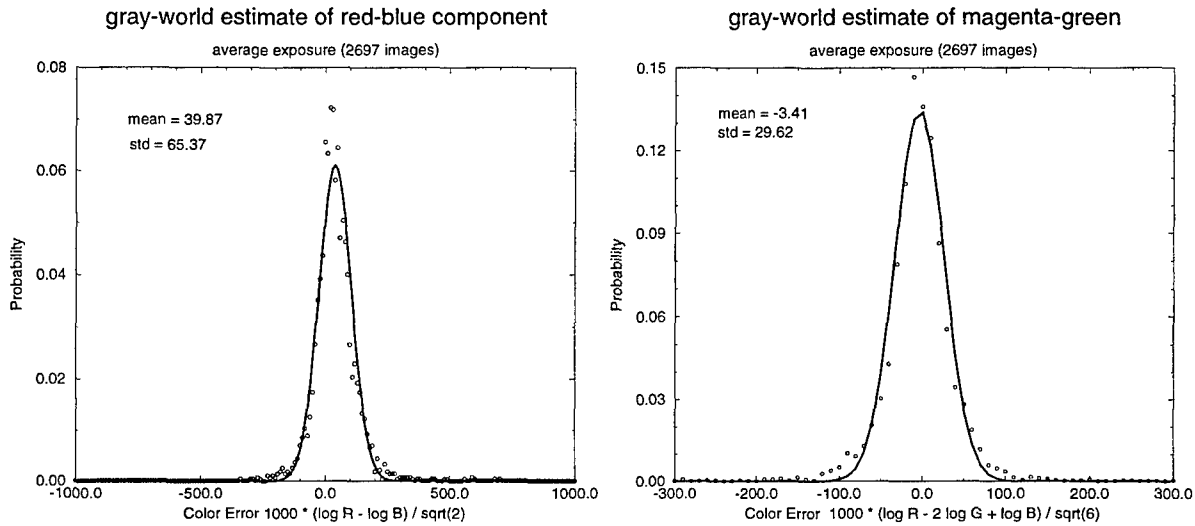

Figure 4. Comparison of red-blue and magenta-green error distributions. Left: error distribution in the red-blue direction; Right: error distribution in the magenta-green direction.

to have higher peaks and wider tails than a Gaussian distribution with the same mean and standard deviation. The gray world estimation of color balance point is clearly much better than its corresponding estimate for the density balance point. The standard deviations from the aim values are much smaller, compared with that shown in Fig. 3.

For detecting skin colors, there are two main approaches. One approach[27–29] is to compile the statistical distribution of skin pixels and use it with other shape and texture cues to decide if a pixel or a region in a new input image belongs to skin color. The other approach is to detect human faces in the image.[30–34] However, as we mentioned before, detecting skin regions does not give us a unique chromaticity pair because skin chromaticities are functions of race, sun tan, scene illumination, and many other factors. Regional, seasonal, and cultural statistics can give us some prior distribution of skin chromaticities to help the algorithms make the best estimates.

Given the neutral and skin colors, we still need two more pairs of chromaticities before we can estimate the matrix M. Other candidate colors are sky, soil, and grass. Unfortunately, their natural chromaticities are even more varied than the skin color. For outdoor scenes, a possible color vector is the daylight locus. It has been shown[35] that for a color imaging system whose spectral response bands are not too wide (say, on the order of 100

nm), the chromaticity distribution of a color image tend to be elongated along the natural daylight locus. This distribution tendency can also be seen in the data reported in other studies.[13] This is mainly due to the mixed illumination of sunlight and skylight on object surfaces. Because the chromaticity distribution of any given color image is heavily biased by the content of the scene, this daylight characteristic can be used only when many images from the same imaging system are available. In practice, this is not an unreasonable constraint because customer orders tend to come in film rolls or image groups.

3.2.3. Image enhancement

Image capture and display/printing processes invariably introduce blur and noise into the images. Image sharpening and noise suppression are two image enhancement operations that have been studied for many years.[36] New algorithms[37–39] using wavelet transforms are also becoming very promising.

In order to sharpen an image and suppress the noise, it is most desirable to have methods for estimating how much and what type of sharpening is needed, and for estimating the noise level as a function of signal in the image. Image blur caused by object motion, focus error, camera optics, film, and scanner can be a complex function to model.[40] In consumer images, image blur is usually not too serious in the sense that most edges are still detectable. An intuitive approach for estimating image blur is to detect all high-contrast, straight edges in the image. By certain heuristic criteria (such as chromatic edges[41] and contrast-normalized gradient[42]), we can locate physical edges that are likely to be straight occlusion edges. The blur function can then be estimated from the edge profiles.[43] Alternatively, edge blur can be modeled and the model parameters estimated from the profiles.[44]

Noise estimation has been studied many times[45,42,44] in the past. A rough estimate of homogeneous, signal-independent white noise is not difficult to compute whenever the image contains some uniform area. However, when the entire image is full of busy textures, all existing methods seem to fail. Fortunately, most consumer images have some uniform area if local shading is removed by polynomial fitting.

4. DEVICES OF UNKNOWN CHARACTERISTICS

In order to achieve good tone and color reproduction, all imaging devices should be carefully calibrated. However, color calibration requires expensive instruments, technical knowledge, and time-consuming efforts. Therefore, most monitors and printers used at home and offices are not calibrated at all. As a consequence, images are typically displayed or printed at less than desirable quality. The chaotic situation is mainly caused by the lack of well accepted standards. The other major contributor is the stability of most imaging devices, whose characteristics change with time, temperature, humidity, usage, and other uncontrollable factors. These two major causes of chaos can be dealt with by consensus of default standards and by development of easy to use tools for characterizing imaging devices either with inexpensive instruments or with visual judgment.

4.1. Default Standards

Standards are driven by competing forces and that is why they are often compromised solutions. However, no standard is worse than a sub-optimum standard. The other driving force is the speed of technology development. It means that trying to perfect a standard may take longer than the life of the current technology.

The ITU-R Recommendation BT.709 forms the basis of many default color standards. Within the international organization ITU, ITU-R is responsible for the coordination for the efficient use of the radio spectrum and of the geostationary satellite orbit.[2] Within this function, it makes recommendations for television broadcasting systems. The basic colorimetric parameters of Recommendation BT.709 for the HDTV standard are as follows.

- The chromaticity coordinates (x,y) of the primaries are:
  red: (0.640,0.330), green: (0.300, 0.600), blue: (0.150, 0.060).
- The white point is D_{65}, (x,y) = (0.3127, 0.3290).

- The overall opto-electronic transfer characteristics at source are:

$$V = 1.099Y^{0.45} - 0.099 \quad \text{for } 0.018 \leq Y \leq 1.0$$
$$V = 4.500Y \quad \text{for } 0.0 \leq Y < 0.018$$

where Y is the relative luminance of the scene and V is the corresponding electrical signal.

If we assume that the video signal is displayed on a CRT monitor with a gamma of 2.22 and a viewing flare of 0.1% of the reference white, then the tone reproduction curve for the HDTV images can be derived. The result is shown in Fig. 5. From the figure, it is obvious that the curve has a slope much higher than one, as required

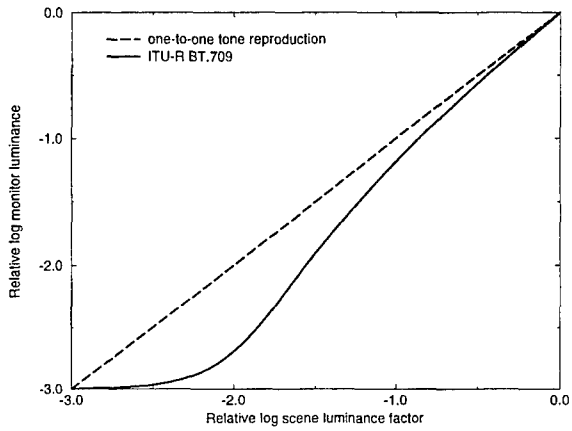

Figure 5. The tone reproduction curve used in the HDTV luminance channel as specified by the international standard (ITU-R BT.709).

by the Bartleson-Breneman's brightness model.[46,47] However, if the viewing flare is more than 0.1%, the actual tone reproduction will not have good contrast in the shadow areas.

Recently, sRGB[48] has become a popular default standard color space, which is based on the same primaries and white point as specified in ITU-R BT.709. Since a typical viewing environment of computer monitors is not in a dark surround as was implied by ITU-R BT.709, the sRGB standard changes the reference viewing environment to a dim surround. The sRGB reference viewing environment is assumed to have a 1% veiling flare, an ambient illuminance level of 64 lux with a D_{50} ambient illuminant, and a proximal field about 20% of the reflectance of the reference display luminance level, which is at 80 cd/m^2. These conditions are specified to facilitate the use of color appearance models (such as CIECAM97) for converting one viewing environment to another.

4.2. Visual Characterization

The characteristics of imaging devices change with time. Some devices (such as monitors) allow users to adjust their settings. Therefore, a printer or a monitor might have been well calibrated in the factory, but over its life time, it cannot consistently reproduce colors well without repeated calibration. Use of ICC device profiles or default color spaces cannot solve this type of problem. What is needed for home users is a convenient way to characterize imaging devices. The best solution is for each device to have a built-in internal self-calibration. The next best solution is to have very inexpensive portable instruments to go with an easy-to-use software tool. Since these two solutions tend to increase the cost of the products, a good alternative solution is to use the user's own eyes as an instrument. Test targets can be displayed or printed with well designed patterns. The user can tell the device driver which pattern is best according to the instructed criteria. The driver then uses the user input to select the current, best calibration table for the device. Several such visual characterization methods have been proposed for printers[49–51] and color monitors.[52,50,53]

There are several perceptual phenomena that can be exploited for the visual characterization of displays and printers. The most frequently used one is visual blur. For example, halftone printing using black dots on white paper can generate an image with fine gray scale shadings indistinguishable from a continuous tone image, if the size of the dot is so small as to be blurred together by the optics of our eyes. The following example shows how visual blur can be used to determine the gamma of a CRT monitor.

The basic idea is to model the input-output characteristics of a CRT monitor by a simple equation with a few parameters, and use visual inspection to select the parameter values by choosing the targets that have the right appearances. For example, the luminance as a function of the input digital value of a CRT can be modeled[54,55] as:

$$L = (S - \delta)^{\gamma} \tag{3}$$

where L is the relative luminance, S is the input digital value, δ is the offset, and γ is the gamma of the channel being considered. Equation (3) does not take external flare into account, and thus is valid only in a completely darkened room. To simplify the example, we will assume that the offset δ has been determined by some other means. We can generate a pattern target that will allow us to determine the correct γ value when it is viewed from a distance. Figure 6 shows a magnified view of the target used for this process. A disk is partitioned into two

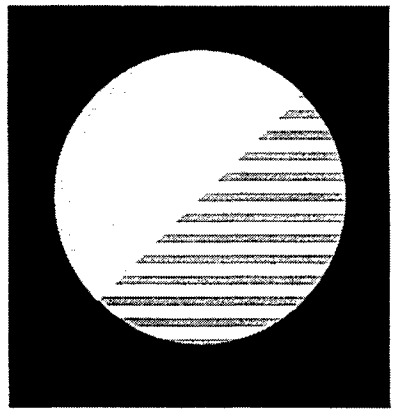

Figure 6. The disk pattern for estimating the CRT gamma.

halves along a 45-degree line. The upper left half is uniformly filled with a single digital value S. The lower right half is filled with alternating dark lines and bright lines. The dark lines have a digital value, S_1, and the bright lines, S_2. If the user is sufficiently far away from the the CRT screen, the dark lines and the bright lines appear to blend together, by the optical blur in the user's eye, to give a shade of gray that is the average luminance of the dark and the bright lines. Two considerations are important for the design of this pattern: (1) The 45 degree boundary is used because our visual system is less sensitive to the oblique direction and therefore can fuse the two sides better when they are of equal luminance. (2) The alternate dark and bright **lines**, instead of a checker board pattern of dark and bright **pixels**, are used because most CRTs cannot display on-off patterns fast enough to produce faithful dark and bright pixels.

Given the offset and the gamma, we can calculate the signal value S on the left half that will match the luminance on the right half:

$$L_1 = (S_1 - \delta)^{\gamma} \tag{4}$$

$$L_2 = (S_2 - \delta)^{\gamma} \tag{5}$$

$$L = (S - \delta)^{\gamma} \tag{6}$$

and

$$L = \frac{1}{2}(L_1 + L_2) \tag{7}$$

$$(S-\delta)^\gamma = \frac{1}{2}[(S_1-\delta)^\gamma + (S_2-\delta)^\gamma]. \tag{8}$$

Therefore,

$$S = (\frac{1}{2}[(S_1-\delta)^\gamma + (S_2-\delta)^\gamma])^{1/\gamma} + \delta \tag{9}$$

To estimate the gamma, we display a series of disks, such as shown in Figure 7, each of which has the same right half with alternating dark and bright lines. The left half is filled with a digital value calculated to match the

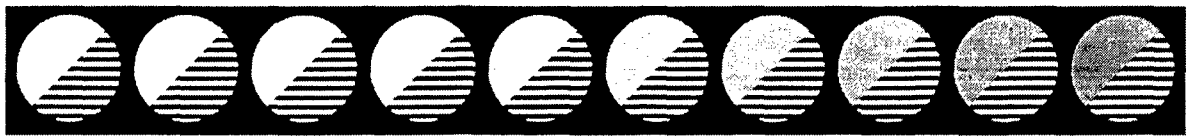

Figure 7. A series of disks for estimating the CRT gamma.

right half, assuming that the CRT has a certain γ value. For example, the first disk is generated with $\gamma = 1.5$, the second with $\gamma = 1.6$, the third with $\gamma = 1.7$, and so on. If the CRT has a γ value of 2.1, then the disk that was generated with $\gamma = 2.1$ will look like a uniform disk with both halves appear to have the same luminance. The user's task is to choose, from the array of disks, the one that seems to have the best match of luminances between the left half and the right half. The chosen disk provides the estimate of the CRT γ, i.e., γ = the gamma value used to generate the selected disk.

A very interesting method conceived by R. L. Gregory for determining the relative "brightness" of different colors is described on page 398 of the book by Kaiser and Boynton.[56] Let a monitor displaying a set of stripes of color A moving to one direction and a set of stripes of color B moving to the opposite direction. Movement is perceived in the direction of the brighter stripes. When both colors are of nearly equal brightness, no drift motion is perceived. Therefore, in principle, it is possible to use this effect to estimate the relative brightness of the red, green, and blue phosphors of a color monitor.

There are many other visual phenomena that are well known in vision research, but have not been well exploited in visual calibration tools. It seems that future research along this direction may produce some solutions to one of the most troublesome problems in Internet color imaging. However, certain visual phenomena are not very sensitive to the variable that we wish to measure. Therefore, search for a robust phenomenon to use is not easy.

5. VIEWING CONDITIONS OF UNKNOWN PERCEPTUAL EFFECTS

The environment in which we view an image has very significant effects on our image perception.[46,3,57] There are three major factors to be considered: (1) visual adaptation, (2) surround effect, and (3) viewing flare. Although color appearance models[57] are developed to predict the effects of such factors, they tend to have many parameters that are not easy to adjust for an arbitrary viewing environment. The best solution for this problem is to set up our viewing environment to one of the standard conditions. However, this is not practical in many applications. In terms of what a user can do, reducing flare by turning or shielding the room illumination away from the monitor or viewing a reflection print from an off-specular angle under a directional light source are common sense actions to take.

If we are producing color images that will be viewed under viewing conditions of unknown perceptual effects, the best strategy is to control the dynamic range of the images by spatial processing[58–64] so that details in both

the highlight and the shadow are preserved with good contrast within a compressed luminance dynamic range. Colors need to be made more saturated and white (or gray) borders or backgrounds can be used to help control the chromatic adaptation of the viewer.

6. DISCUSSION AND CONCLUSIONS

Standardization across all imaging devices is the main solution to the problems of Internet color imaging. However, standardization does not solve all the problems. The three remaining problems, as discussed in this paper, are quite different in nature and require different types of solution. To deal with color images of unknown calibration, research in computer vision, image understanding, and scene physics will eventually allow us to implement automatic algorithms to handle the problem. To deal with imaging devices of unknown characteristics, inexpensive colorimeters and easy-to-use software calibration tools will be the most feasible solutions in the near future. To deal with viewing conditions of unknown perceptual effects, users can take simple measures to greatly improve their image perception. The alternative solution is to build display devices that can sense the environments and self-adjust their own tone and color reproduction characteristics.

ACKNOWLEDGMENTS

Many of the results reported here are from the joint work with my colleagues. I would like to thank them and gladly acknowledge their contributions: James Alkofer, John Birkelund, Scott Daly, Robert Goodwin, Heemin Kwon, and Jeanie Liang.

REFERENCES

1. D. McDowell, "The role and responsibilities of ISO/IEC joint technical advisory group 2 - imagery," *Proc. CIE Expert Symposium '96 on Colour Standards for Image Technology*, pp. 4–6, 1996.
2. C. J. Dalton, K. P. Davies, and O. Gofaizen, "An overview of the activities of the itu in television colorimetry," *Proc. CIE Expert Symposium '96 on Colour Standards for Image Technology*, pp. 18–24, 1996.
3. E. J. Giorgianni and T. E. Madden, *Digital Color Management*, Addison-Wesley, Reading, MA., 1997.
4. M. Stokes and R. Motta, "A default RGB monitor space proposal," *Proc. CIE Expert Symposium '96 on Colour Standards for Image Technology*, pp. 47–50, 1996.
5. M. Stokes, "Color management in the real world: sRGB, ICM2, ICC, ColorSync^{TM}, and other attempts to make color management 'transparent'," *Proc. SPIE* **3299**, pp. 360–367, 1998.
6. L. MacDonald, "Colour management and display calibration," *Proc. CIE Expert Symposium '96 on Colour Standards for Image Technology*, pp. 63–69, 1996.
7. A. Johnson, *Colour Management in Graphic Arts and Publishing*, Pira International, Leatherhead, Surrey, UK., 1996.
8. F. Deleixhe-Mauhin, J. M. Krezinski, G. Rorive, and G. E. Pierard, "Quantification of skin color in patients undergoing maintenance hemodialysis," *J. Am. Acad. Dermatol.* **27, 6, Part 1**, pp. 950–953, 1992.
9. G. E. Pierard, C. Pierard-Franchimont, F. L. Dosal, T. B. Mosbah, J. Arrese-Estrada, A. Rurangirwa, A. Dowlati, and M. Vardar, "Pigmentary changes in skin senescence," *J. Appl. Cosmetol.* **9**, pp. 57–63, 1991.
10. Y. Yamamoto, "Colorimetric evaluation of skin color in the Japanese," *Plast. Reconstr. Surg.* **96, 1**, pp. 139–145, 1995.
11. E. Edwards and S. Duntley, "Pigment and color in living human skin," *Am. J. Anat.* **65**, pp. 1–33, 1939.
12. E. Edwards and S. Duntley, "Analysis of skin pigment changes after exposure to sunlight," *Science* **90**, pp. 235–237, 1939.
13. G. Burton and I. Moorhead, "Color and spatial structure in natural scenes," *Appl. Opt.* **26, 1**, pp. 157–170, 1987.
14. D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Am.* **A, 4, 12**, pp. 2379–2394, 1987.
15. D. L. Ruderman, "Origins of scaling in natural images," *Vision Res.* **37, 23**, pp. 3385–3395, 1997.
16. J. S. Alkofer, *Tone Value Sample Selection in Digital Image Processing Method Employing Histogram Normalization*, U.S. Patent No. 4,654,722, Mar. 31, 1987.

17. W. Richards, "Lightness scale from image intensity distribution," *Appl. Opt.* **21, 14**, pp. 2569–2582, 1982.
18. J. Huang and D. Mumford, "Statistics of natural images and models," *Proc. IEEE Conf. Comput. Vision Pattern Recognit.* **1**, pp. 541–547, 1999.
19. H.-C. Lee and H. Kwon, *Method for Estimating and Adjusting Digital Image Contrast*, U.S. Patent No. 5,822,453, Oct. 13, 1998.
20. L. A. Jones and H. R. Condit, "The brightness scale of exterior scenes and the computation of correct photographic exposure," *J. Opt. Soc. Am.* **31, 11**, pp. 651–678, 1941.
21. J. S. Alkofer, *Contrast Adjustment in Digital Image Processing Method Employing Histogram Normalization*, U.S. Patent No. 4,731,671, Mar. 15, 1988.
22. R. M. Evans, *Method for Correcting Photographic Color Prints*, U.S. Patent No. 2,571,697, Oct. 16, 1951.
23. J. Hughes and J. K. Bowker, "Automatic color printing techniques," *Image Technol.*, pp. 39–43, April/May 1969.
24. H.-C. Lee, L. L. Barski, and R. A. Senn, *Automatic Tone Scale Adjustment Using Image Activity Measures*, U.S. Patent No. 5,633,511, May 27, 1997.
25. J. R. Boyack and A. K. Juenger, *Brightness Adjustment of Images Using Digital Scene Analysis*, U.S. Patent No. 5,724,456, Mar. 3, 1998.
26. H.-C. Lee and R. Goodwin, "Colors as seen by humans and machines," *Final Program and Advance Printing Papers of the IS&T's 47th Annual Conference*, pp. 401–405, 1994.
27. Y. Satoh, Y. Miyake, H. Yaguchi, and S. Shinohara, "Facial pattern detection and color correction from negative color film," *J. Imaging Technol.* **16, 2**, pp. 80–84, 1990.
28. D. A. Forsyth and M. M. Fleck, "Automatic detection of human nudes," *Int. J. Comput. Vision* **32, 11**, pp. 63–77, 1999.
29. M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection," *Proc. IEEE Conf. Comput. Vision Pattern Recognit.* **1**, pp. 275–281, 1999.
30. M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognitive Neurosci.* **3, 1**, pp. 71–86, 1991.
31. G. Yang and T. Huang, "Human face detection in complex background," *Pattern Recognit.* **27, 1**, pp. 53–63, 1994.
32. K. Yow and R. Cipolla, "Feature-based human face detection," *Image Vision Comput.* **15**, pp. 713–735, 1997.
33. K. K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **20, 1**, pp. 39–51, 1998.
34. H. Rowley, S. Baluja, and T. Kanade, "Neural-network-based face detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **20, 1**, pp. 23–38, 1998.
35. H.-C. Lee, "A physics-based color encoding model for images of natural scenes," *Proc. Conf. Mod. Eng. Technol., Electro-Optics Session*, pp. 25–52, 1992.
36. B. Bayer and P. Powell, "A method for the digital enhancement of unsharp, grainy photographic images," *Adv. Comput. Vision Image Proc.* **2**, pp. 31–88, 1986.
37. D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Am. Statistical Association* **90, 432**, pp. 1200–1224, 1995.
38. J. Lu, D. M. Healy, Jr., and J. Weaver, "Contrast enhancement of medical images using multiscale edge representation," *Opt. Eng.* **33, 7**, pp. 2151–2161, 1994.
39. E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms," *IEEE Trans. Inf. Theory* **38, 2**, pp. 587–606, 1992.
40. H.-C. Lee, "A review of of image-blur models in a photographic system using the principles of optics," *Opt. Eng.* **29, 5**, pp. 405–421, 1990.
41. H.-C. Lee, "Chromatic edge detection: Idealization and reality," *Int. J. Imaging Syst. Technol.* **2**, pp. 251–266, 1990.
42. J. Katajamäki and H. Saarelma, "Objective quality potential measures of natural color images," *J. Imaging Sci. Technol.* **42, 3**, pp. 250–263, 1998.
43. S. E. Reichenbach, S. K. Park, and R. Narayanswamy, "Characterizing digital image acquisition devices," *Opt. Eng.* **30, 2**, pp. 170–177, 1991.

44. V. Kayargadde and J. B. Martens, "Estimation of edge parameters and image blur using polynomial transforms," *CVGIP: Graphic Models and Image Processing* **56, 6**, pp. 442–461, 1994.
45. S. I. Olsen, "Estimation of noise in images: an evaluation," *CVGIP: Graphic Models and Image Processing* **55, 4**, pp. 319–323, 1993.
46. A. Johnson and J. Birkenshaw, "The influence of viewing conditions on colour reproduction objectives," *Proc. of the 14th International Conference of Printing Research Institutes*, pp. 48–72, 1977.
47. H.-C. Lee, S. Daly, and R. L. Van Metter, "Visual optimization of radiographic tone scale," *Proc. SPIE* **3036**, pp. 118–129, 1997.
48. *IEC 61966. Part 2.1: Default RGB colour space - sRGB (Third working draft)*, International Electrotechnical Commission, 1998.
49. S. J. Harrington, *Printer Calibration Using a Tone Reproduction Curve and Requiring No Measuring Equipment*, U.S. Patent No. 5,347,369, Sep. 13, 1994.
50. A. D. Edgar and J. M. Kasson, *Display Calibration*, U.S. Patent No. 5,298,993, Mar. 29, 1994.
51. K. A. Hadley and K. E. Spaulding, *Method for Printer Calibration*, U.S. Patent No. 5,995,714, Nov. 30, 1999.
52. R. J. Motta, "Visual characterization of color CRTs," *Proc. SPIE* **1909**, pp. 212–221, 1993.
53. S. J. Daly and H.-C. Lee, *Visual Characterization Using Display Model*, U.S. Patent No. 5,754,222, May 19, 1998.
54. R. Bartow, W. Darrow, and T. Hartmann, *CRT Device Light Versus Input Signal Characteristic Function*, U.S. Patent No. 4,862,265, Aug. 29, 1989.
55. R. Berns, R. Motta, and M. Gorzynski, "CRT colorimetry. Part I: Theory and practice," *Color Res. Appl.* **18, 5**, pp. 299–314, 1993.
56. P. Kaiser and R. Boynton, *Human Color Vision*, Optical Society of America, Washington, D.C., second ed., 1996.
57. M. Fairchild, *Color Appearance Models*, Addison-Wesley, Reading, MA., 1997.
58. E. Wagensonner, W. Ruf, H. Fuchsberger, and K. Birgmeir, *Method of Electronically Improving the Sharpness and Contrast of a Colored Image for Copying*, U.S. Patent No. 4,812,903, Mar. 14, 1989.
59. H.-C. Lee, M. Kaplan, and R. Goodwin, *An Interactive Dynamic Range Adjustment System for Printing Digital Images*, U.S. Patent No. 5,012,333, Apr. 30, 1991.
60. P. Vuylsteke and E. Schoeters, *Method and Apparatus for Contrast Enhancement*, U.S. Patent No. 5,467,404, Nov. 14, 1995.
61. M. Nakazawa and H. Tsuchino, *Method for Compressing a Dynamic Range for a Radiation Image*, U.S. Patent No. 5,471,987, Dec. 5, 1995.
62. H. Tsuchino and M. Nakazawa, *Radiation Image Processing Method Which Increases and Decreases a Frequency Region of the Radiation Image*, U.S. Patent No. 5,493,622, Feb. 20, 1996.
63. N. Nakajima, *Method for Compressing Dynamic Ranges of Images Using a Monotonously Decreasing Function*, U.S. Patent No. 5,608,813, Mar. 4, 1997.
64. F. Labaere and P. Vuylsteke, *Image Contrast Enhancing Method*, U.S. Patent No. 5,717,791, Feb. 10, 1998.